

Serial No. 10/059,421  
Docket No. YOR920010126US2

2

These rejections are respectfully traversed in the following discussion.

## **I. THE CLAIMED INVENTION**

The claimed invention (e.g., as recited in claim 1) is directed to a system for identifying genes. The system includes a pattern database comprising patterns of amino acids, and an input device for inputting a genomic DNA sequence.

Importantly, the system further includes a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database to determine whether the open reading frame includes a putative gene in the DNA sequence.

Conventional systems for identifying genes (e.g., putative genes) are either based on the use the statistics of DNA sequences, or the use of similarity searches to determine gene locations (Application at page 2, lines 7-22). However, these conventional methods have various problems which prevent them from efficiently identifying genes in a given DNA sequence (Application at page 3, line 19-page 4, line 21).

The claimed invention, on the other hand, includes a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database to determine whether the open reading frame includes a putative gene in the DNA sequence (Application at Figure 1; page 5, lines 4-11). The claimed invention may be considered as including the best characteristics of statistical approaches and database similarity searches, in identifying genes in a given DNA sequence (Application at page 6, lines 18-21).

## **II. THE ALLEGED PRIOR ART REFERENCES**

### **A. Rigoutsos '99**

The Examiner alleges that Rigoutsos '99 teaches the claimed invention of claims 1-8, 10, 11, 13-23, 25 and 27-29. Applicant submits, however, that there are features of the claimed

Serial No. 10/059,421  
Docket No. YOR920010126US2

3

invention which are neither taught nor suggested by Rigoutsos '99.

Rigoutsos '99 discloses a process for building dictionaries of motifs from amino acid sequences. Attached hereto as Exhibit 1 is a flowchart illustrating the process of Rigoutsos '99.

Applicant respectfully submits that contrary to the Examiner's allegations, Rigoutsos '99 does not teach or suggest *"a processor which: translates an open reading frame (ORF) of said DNA sequence into an amino acid translation; and locates in said amino acid translation occurrences of said patterns from said pattern database to determine whether said open reading frame includes a putative gene in said DNA sequence"*, as recited, for example, in claim 1, and similarly recited in claims 16, 23 and 29.

As noted above, unlike conventional systems for identifying genes (e.g., putative genes) which are either based on the use of the statistics of DNA sequences, or the use of similarity searches to determine gene locations, the claimed invention includes a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database to determine whether the open reading frame includes a putative gene in the DNA sequence (Application at Figure 1; page 5, lines 4-11). The claimed invention may be considered as including the best characteristics of statistical approaches and database similarity searches, in identifying genes in a given DNA sequence (Application at page 6, lines 18-21).

Clearly, these features are not taught or suggested by the Rigoutsos '99. Indeed, Applicant would reiterate the points that Applicant made in the Supplemental Amendment filed herein on June 22, 2006:

**1. Rigoutsos '99 does not teach or suggest translating an ORF of the input DNA sequence into an amino acid translation**

Applicant again submits that the Examiner's position here is unclear. The Examiner first indicates on page 3 of the Office Action that this feature is disclosed at page 224, col. 1, lines 6-40. However, later on page 3 the Examiner concedes that Rigoutsos '99 "does not explicitly

Serial No. 10/059,421  
Docket No. YOR920010126US2

4

describe the limitation of "translates an open reading frame", but implies that this feature is inherent from the disclosure on page 225, column 2.

However, Applicant would submit that nowhere does col. 1 on page 224 or col. 2 (the "Database" Section) on page 225 teach or suggest translating an ORF of an input DNA sequence into an amino acid translation. Indeed, col. 1 on page 224 merely states that 1) a database containing the ORFs for the complete genomes of 13 bacteria were compiled, 2) the database was processed to produce a dictionary of seqlets, and 3) the seqlets were aligned in three-dimensional space. Further, as noted above, the "Database" Section on page 225 merely states that:

*"[t]he database that we used as the input to Teiresias comprised the ORFs from the 17 complete and publicly available archaeal and bacterial genomes. In particular, the database contained: 4289 ORFs from Escherichia coli with 1358990 [amino acids] ..."*

Nowhere in these passages does Rigoutsos '99 even mention translating an ORF of an input DNA sequence into an amino acid translation. In fact, the Examiner seems to imply (in paragraph 9 on pages 3-4 of the Office Action) that Rigoutsos '99 discloses amino acid sequences, and that somebody had to have translated these amino acid sequences from a DNA sequence at some point, then Rigoutsos '99 discloses translating a DNA sequence into an amino acid translation.

However, the Examiner is missing an important point. It is not any old DNA sequence that is being translated in the claimed invention. Instead, the claim recites **translating an ORF of the input DNA sequence into an amino acid translation**. That is, it is an ORF of the input (e.g., query) DNA sequence under examination that is being translated.

That is, even assuming (arguendo) that the amino acid sequences in Rigoutsos '99 were at some point in time translated by some unknown entity from a DNA sequence, nowhere does Rigoutsos '99 teach or suggest a system which inputs a DNA sequences and translates an ORF of

Serial No. 10/059,421  
Docket No. YOR920010126US2

5

that input sequence into an amino acid sequence.

**2. Rigoutsos '99 does not teach or suggest locating in the amino acid translation occurrences of patterns from the pattern database to determine whether the ORF includes a putative gene**

Applicant would point out that this element of the claimed invention may be considered in two parts:

1) locating occurrences of patterns from the pattern database in the amino acid translation; and

2) determining (e.g., based on whether the patterns occur in the amino acid translation) whether the ORF includes a putative gene.

The Examiner alleges in paragraph 8 on page 3 of the Office Action that this feature of the claimed invention is taught at page 228, col. 2, lines 4-11 in Rigoutsos '99. However, this passage merely reads (in pertinent part)

*"[o]ne of the seqlets that are discovered when we process the input database is .... and is present in the following ten ORFs: ..... Of these, gi\_3328856 and gi\_3329230 are annotated as Fe-S oxidoreductases."*

That is, this simple passage clearly does not teach or suggest 1) locating occurrences of patterns from the pattern database in the amino acid translation, or 2) determining whether the ORF includes a putative gene.

In fact, when asked at the interview conducted on June 14, 2006, where in this passage these features are disclosed, the Examiner stated that the features were disclosed "at other places" in Rigoutsos '99. Applicant respectfully requested that, in a next office action, the Examiner more clearly indicate the specific disclosure in Rigoutsos '99 on which the Examiner was relying as disclosing these features of the claimed invention.

However, Applicant would point out that nowhere in this Office Action has the Examiner indicated such specific disclosure. Instead, the Examiner again simply relies on page 228, col. 2,

Serial No. 10/059,421  
Docket No. YOR920010126US2

6

lines 4-11 in Rigoutsos '99. Therefore, Applicant again requests that the Examiner **more clearly indicate the specific disclosure in Rigoutsos '99 on which the Examiner was relying as disclosing these features of the claimed invention.**

Indeed, Applicant would again point out that the passage at page 228, col. 2, lines 4-11 in Rigoutsos '99 clearly does not teach or suggest locating occurrences of patterns from the pattern database in the amino acid translation. Indeed, nowhere does this passage discuss "locating" anything, let alone locating occurrences of "patterns" in an amino acid translation (from the input DNA sequence).

Further, the Examiner seems to misunderstand the term "putative gene". Indeed, the Examiner seems to imply that this passage in Rigoutsos '99 teaches identifying a putative gene from the mere disclosure that two amino acid sequences (gi\_3328856 and gi\_3329230) are annotated as oxidoreductases. This is clearly unreasonable.

Indeed, an oxidoreductase is simply an enzyme that catalyzes the transfer of electrons from one molecule to another. An enzyme may be but is not necessarily a protein.

Thus, it is clear that Rigoutsos '99 merely discloses in this passage that two amino acid sequences are annotated as a certain enzyme. Nowhere does Rigoutsos '99 describe how, where or why these amino acids were so annotated, Rigoutsos '99 simply states that the sequences were so annotated. **Clearly, this passage does not teach or suggest determining whether an ORF on an input DNA sequence includes these two amino acid sequences.**

Therefore, nowhere in this passage or anywhere else, does Rigoutsos '99 teach or suggest 1) locating occurrences of patterns from the pattern database in the amino acid translation, and 2) determining (e.g., based on whether the patterns occur in the amino acid translation) whether the ORF includes a putative gene.

**3. Rigoutsos '99 does not teach or suggest a "system for identifying genes" or "inputting a genomic DNA sequence"**

Applicant again notes that claim 1 recites a "system for identifying genes". However,

Serial No. 10/059,421  
Docket No. YOR920010126US2

7

Rigoutsos '99 has nothing to do with "identifying genes". Indeed, this is clear from the title of Rigoutsos '99 which is "***Building Dictionaries Of 1D and 3D Motifs By Mining The Unaligned 1D Sequences of ... Genomes***". That is, Rigoutsos '99 deals with building dictionaries of motifs (i.e., from amino acid sequences), not identifying genes.

In fact, Rigoutsos '99 simply discloses a process that may be used to create a pattern database. Indeed, Rigoutsos '99 is summarized on page 224, col. 1 which states "*[w]e report on the properties of the entries of these two dictionaries, the extent to which the 1D seqlets have instances in the Protein Data Bank, and the ramifications from the induced coverage. We also present several entries of the generated dictionaries and discuss uses*". As is clear from this summary, Rigoutsos '99 has nothing to do with 1) inputting a genomic DNA sequence, 2) translating an ORF of the DNA sequence into an amino acid translation, 3) locating in the amino acid translation occurrences of patterns from the pattern database, or 4) determining whether the ORF includes a putative gene.

In fact, as clearly indicated in the flowchart illustrated in the attached Exhibit 1, which, Rigoutsos '99 has nothing to do with identifying genes and clearly does not teach or suggest the claimed invention.

Further, with respect to the limitation "**inputting a genomic DNA sequence**", the Examiner relies on two simple words in the Abstract to support his position that Rigoutsos '99 teaches a DNA sequence. Specifically, the Examiner alleges that the phrase "genomic input" in line 6 of the Abstract necessarily implies DNA.

However, Applicant would point out that the term "genomic input" is modified by the language following it. That is, the complete term used in the Abstract is "**genomic input at the level of amino acid positions**", which makes clear that the input in Rigoutsos '99 is not a DNA sequence (indeed, the purpose of the Rigoutsos '99 paper is to build a dictionary of motifs from amino acid sequences not from DNA sequences).

Moreover, even assuming (arguendo) that the Abstract may be construed as disclosing a DNA sequence, the remainder of Rigoutsos '99 makes it abundantly clear that only amino acid

Serial No. 10/059,421  
Docket No. YOR920010126US2

8

sequences are being input, not DNA sequences. Indeed, for example, on page 225, Rigoutsos '99 states that

*"[t]he database that we used as the input to Teiresias comprised the ORFs from the 17 complete and publicly available archaeal and bacterial genomes. In particular, the database contained: 4289 ORFs from Escherichia coli with 1358990 [amino acids] ..."*

Indeed, nowhere in Rigoutsos '99 is the term DNA even used. Instead, Rigoutsos '99 only refers to amino acid sequences. Thus, it is clear that contrary to the Examiner's allegations Rigoutsos '99 does not teach or suggest inputting a DNA sequence.

Therefore, Applicant submits that there are features of the claimed invention that are not taught or suggested by Rigoutsos. Therefore, the Examiner is respectfully requested to withdraw this rejection.

#### **B. Delcher**

The Examiner alleges that Rigoutsos would have been combined with Delcher to form the invention of claims 9 and 26. Applicant submits, however, that Rigoutsos would not have been combined with Delcher and even if combined, the combination would not teach or suggest each and every feature of the claimed invention.

Applicant respectfully submits that these references would not have been combined as alleged by the Examiner. Indeed, these references are unrelated, and no person of ordinary skill in the art would have considered combining these disparate references, absent impermissible hindsight.

In fact, these references clearly do not teach or suggest their combination. Therefore, Applicant respectfully submits that one of ordinary skill in the art would not have been so motivated to combine the references as alleged by the Examiner. Therefore, the Examiner has failed to make a prima facie case of obviousness.

Moreover, contrary to the Examiner's allegations, neither Rigoutsos nor Delcher, nor any

Serial No. 10/059,421  
Docket No. YOR920010126US2

9

combination thereof teaches or suggests *"a processor which: translates an open reading frame (ORF) of said DNA sequence into an amino acid translation; and locates in said amino acid translation occurrences of said patterns from said pattern database to determine whether said open reading frame includes a putative gene in said DNA sequence"*, as recited, for example, in claim 1, and similarly recited in claims 16, 23 and 29.

As noted above, unlike conventional systems for identifying genes (e.g., putative genes) which are either based on the use of the statistics of DNA sequences, or the use of similarity searches to determine gene locations, the claimed invention includes a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database to determine whether the open reading frame includes a putative gene in the DNA sequence (Application at Figure 1; page 5, lines 4-11). The claimed invention may be considered as including the best characteristics of statistical approaches and database similarity searches, in identifying genes in a given DNA sequence (Application at page 6, lines 18-21).

Clearly, Delcher does not teach or suggest these features of the claimed invention. Indeed, Applicant would again point out that Delcher does not apply at least because of:

- 1) Delcher uses nucleotide sequences for training;
- 2) Delcher uses sequences from archaeal and bacterial genomes only for training;
- 3) Delcher builds a Hidden Markov Model (HMM) from carefully-selected sequences;
- 4) Delcher uses the HMM to query nucleic acid sequences and compare the expectation in nucleotide sequence space as captured by the HMM with what is encountered in the input at hand;
- 5) In Delcher the comparison is done in the space of nucleotide sequences;
- 6) In Delcher if its nucleotide sequences matches the expectation in terms of nucleotide sequence composition as the latter is captured by the HMM, then an ORF is reported as a putative gene.



Serial No. 10/059,421  
Docket No. YOR920010126US2

10

On the other hand, the claimed invention may be completely different and unrelated to Delcher at least because of:

- 1) the claimed invention may begin with the largest possible database of amino acid sequences;
- 2) the claimed invention may use sequences from archaeal, bacterial, eukaryotic and viral genomes;
- 3) the claimed invention may convert the knowledge captured by the collection of amino acid sequences to a different representation, one of patterns  $p$  derived through unsupervised pattern discovery on these amino acid sequences;
- 4) the claimed invention may translate a given ORF into a putative amino acid sequence  $S$ ;
- 5) the claimed invention may examine  $S$  to determine whether it contains one or more patterns  $p$  from the collection that was discovered above (e.g., in step 3);
- 6) the claimed invention may use the number of patterns  $p$  that have instances in  $S$  to decide whether the putative amino acid sequence is well-formed;
- 7) the claimed invention may carry out comparisons in the space of amino acid sequences;
- 8) in the claimed invention, if the putative amino acid sequence is well-formed then the ORF from which it was generated is reported as a putative gene;

Therefore, contrary to the Examiner's allegations, neither Rigoutsos nor Delcher nor any alleged combination thereof teaches or suggests a processor which translates an open reading frame (ORF) of the DNA sequence into an amino acid translation, and locates in the amino acid translation occurrences of the patterns from the pattern database to determine whether the open reading frame includes a putative gene in the DNA sequence, as in the claimed invention.

Therefore, Applicant submits that these alleged references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the

Serial No. 10/059,421  
Docket No. YOR920010126US2

11

claimed invention. Therefore, the Examiner is respectfully requested to withdraw this rejection.

### III. FORMAL MATTERS AND CONCLUSION

In view of the foregoing, Applicant submits that claims 1-30, all the claims presently pending in the application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Assignee's Deposit Account No. 50-0510.

Respectfully Submitted,

Date: 8/28/06



Phillip E. Miller, Esq.  
Registration No. 46,060

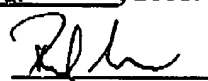
**McGinn IP Law Group, PLLC**  
8321 Old Courthouse Road, Suite 200  
Vienna, VA 22182-3817  
(703) 761-4100  
Customer No. 21254

Serial No. 10/059,421  
Docket No. YOR920010126US2

12

**CERTIFICATE OF FACSIMILE TRANSMISSION**

I hereby certify that the foregoing Amendment was filed by facsimile with the United States Patent and Trademark Office, Examiner C. Dune Ly, Group Art Unit # 2168 at fax number 571-272-8300 this 28th day of August, 2006.

  
\_\_\_\_\_  
Phillip E. Miller  
Reg. No. 46,060